(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number : **0 585 083 A1**

(12) **EUROPEAN PATENT APPLICATION**

(21) Application number : **93306628.4**

(51) Int. Cl.[5] : **C12N 15/54**, C12N 9/10

(22) Date of filing : **20.08.93**

(30) Priority : **21.08.92 JP 243984/92**

(43) Date of publication of application :
02.03.94 Bulletin 94/09

(84) Designated Contracting States :
DE FR GB

(71) Applicant : **TAKARA SHUZO CO. LTD.**
609 Takenakacho
Fushimi-ku Kyoto (JP)

(72) Inventor : **Nishikawa, Atsushi**
3-8-A5-401, Shinsenri Minamimachi
Toyonaka-shi, Osaka-fu (JP)
Inventor : **Taniguchi, Naoyuki**
2-19-32-201, Ueno Higashi
Toyonaka-shi, Osaka-fu (JP)

(74) Representative : **Wakerley, Helen Rachael**
Reddie & Grose, 16 Theobalds Road
London WC1X 8PL (GB)

(54) **Human glycosyltransferase gene.**

(57) A human GnT-III gene having a specific restriction map represented by Fig. 1 of the attached Drawings and a length of approximately 1.6 kb is disclosed, together with a human GnT-III gene which is hybridizable with said gene. A process for producing human GnT-III is provided, which comprises incubating a transformant wherein one of the above-mentioned genes is used and harvesting. The sequencing listing shows a part of the base sequence.

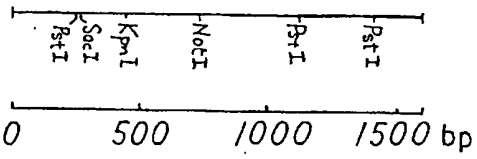

Figure 1

EP 0 585 083 A1

and examining the capability of the transformant to express human GnT-III. Examples of usable expression cells includ COS-1 cells (ATCC CRL 1650). For example, the COS-1 cells can be transformed by the above expression plasmid EX20F. Then the transformant is incubated and the activity of GnT-III expressed in the transformant is determined to specify a gene coding for human GnT-III. This gene is integrated into EX20F and a part of its base sequence is located on a DNA fragment represented by SEQ ID NO. 1. Human GnT-III can be produced by genetic engineering technique by incubating the above transformant.

By effecting hybridization with the use of the gene thus obtained as a probe under strict conditions, it is anticipated that genes for enzymes analogous to that of the present invention, which are different therefrom in sequence but expected to have a similar activity, may be obtained. The term "under stringent conditions" as used herein means that the hybridization of a nylon membrane having DNAs immobilized thereon with the probe is conducted in a solution containing 6 x SSC (1 x SSC means a solution prepared by dissolving 8.76 g of sodium chloride and 4.41 g of sodium citrate in 1 liter of water), 1% of sodium lauryl sulfate, 100 μg/ml of salmon sperm DNA, and 5 x Denhardt's (containing bovine serum albumin, polyvinylpyrrolidone and Ficoll each at a concentration of 0.1%) at 65°C for 20 hours.

As described above in detail, the present invention enables a gene coding for human GnT-III to be isolated and provides a process for producing human GnT-III by using the gene. This gene and its decomposition products are usable in the determination of human GnT-III during the expression process in vivo and, therefore, useful in the genetic diagnosis of cancer, and so forth. In addition, various antibodies can be immunologically prepared by using polypeptides coded for by the gene of the present invention. These antibodies are also useful in the field of diagnosis and for the purification of human GnT-III.

Brief Description of the Drawings

Figure 1 is a drawing showing a restriction map of a gene coding for human GnT-III. Figure 2 is a drawing showing relationships among four DNAs H2, H3, H15 and H20. Figure 3 is a drawing showing the construction of a plasmid EX20F.

Example

To further illistrates the present invention in greater detail, and not by way of limitation, the following Examples will be given.

Example 1

(1) Screening of cDNA library

SV3 was prepared from Escherichia coli XL1-Blue SV3 (FERM BP-4325) transformed by a plasmid SV3 and the plasmid was digested with HindIII to give a DNA fragment of approximately 1.4 kb. This DNA fragment was radiolabeled with [α-^{32}p] dCTP (3000 Ci/mmol, Amersham) by using a Multiprime DNA Labeling System (Amersham) to thereby give a probe. By using the obtained probe, a human cDNA library [Human Fetal Liver <λgt10>, Clonetech] was screened for the target clone by plaque hybridization. As a result, two positive clones were obtained from 3 x 10^6 plaques. From these clones, DNAs were extracted and digested with EcoRI. The digestion products thus obtained were subcloned into Bluescript IISK$^+$ and the DNAs thus subcloned (approximately 1.3 kb and approximately 1.5 kb) were respectively named H2 and H3, while the plasmids were respectively named pBluescript II (H2) and pBluescript II (H3). Figure 2 shows the restriction maps of these DNAs and a relationship between them.

(2) Cloning of upstream region containing initiator codon

H2 and H3 were radiolabeled in the same manner as the one described in the above Example 1-(1) to thereby give probes. By using these probes, screening of a human cDNA library was carried out in the same manner as the one described in the above Example 1-(1) to obtain four positive clones from 7 x 10^5 plaques. The EcoRI-digestion products thereof were subcloned into Bluescript IISK$^+$. The base sequences of the DNAs thus subcloned were identified and two DNAs containing an initiator codon (approximately 1.6 kb and approximately 1.5 kb) were named respectively H15 and H20, while the plasmids corresponding thereto were named respectively pBluescript II (H15) and pBluescript II (H20). Figure 2 shows the restriction maps of these DNAs and a relationship between thereof with H2 and H3.

Example 2

(1) Construction of expression plasmid

Sequence Listing

SEQ ID NO:1

LENGTH:2247

TYPE:nucleic acid

STRANDEDNESS:double

TOPOLOGY:linear

MOLECULE TYPE:cDNA to mRNA

SEQUENCE DESCRIPTION:SEQ ID NO:1:

```
CCGGCTGCGA TGCCGGGCGC CCGCCGCAGC CGCTGCCGCC GGAGCCCGGG ATGGGGCGAG    60
AGGCTGCGGC GGACGCCAGC ATCTCCCCGC CGGGGACCCC GGGGGCCGCG GAGCCGCCGC   120
CGCCGCTGCT GCCGCCGTTG CTGAGACCCA GCGGGCGATG GGATGAAGAT GAGACGCTAC   180
AAGCTCTTTC TCATGTTCTG TATGGCCGGC CTGTGCCTCA TCTCCTTCCT GCACTTCTTC   240
AAGACCCTGT CCTATGTCAC CTTCCCCCGA GAACTGGCCT CCCTCAGCCC TAACCTGGTG   300
TCCAGCTTTT TCTGGAACAA TGCCCCGGTC ACGCCCCAGG CCAGCCCCGA GCCAGGAGGC   360
CCTGACCTGC TGCGTACCCC ACTCTACTCC CACTCGCCCC TGCTGCAGCC GCTGCCGCCC   420
AGCAAGGCGG CCGAGGAGCT CCACCGGGTG GACTTGGTGC TGCCCGAGGA CACCACCGAG   480
TATTTCGTGC GCACCAAGGC CGGCGGCGTC TGCTTCAAAC CCGGCACCAA GATGCTGGAG   540
AGGCCGCCCC CGGGACGGCC GGAGGAGAAG CCTGAGGGGG CCAACGGCTC CTCGGCCCGG   600
CGGCCACCCC GGTACCTCCT GAGCGCCCGG GAGCGCACGG GGGGCCGAGG CGCCCGGCGC   660
AAGTGGGTGG AGTGCGTGTG CCTGCCCCGC TGGCACGGAC CCAGCTGCGG CGTGCCCACT   720
GTGGTGCAGT ACTCCAACCT GCCCACCAAG GAGCGGCTGG TGCCCAGGGA GGTGCCGCGC   780
CGCGTCATCA ACGCCATCAA CGTCAACCAC GAGTTCGACC TGCTGGACGT GCGCTTCCAC   840
GAGCTGGGCG ACGTGGTGGA CGCCTTTGTG GTGTGCGAGT CCAACTTCAC GGCTTATGGG   900
GAGCCGCGGC CGCTCAAGTT CCGGGAGATG CTGACCAATG GCACCTTCGA GTACATCCGC   960
CACAAGGTGC TCTATGTCTT CCTGGACCAC TTCCCGCCCG GCGGCCGGCA GGACGGCTGG  1020
ATCGCCGACG ACTACCTGCG CACCTTCCTC ACCCAGGACG GCGTCTCGCG GCTGCGCAAC  1080
CTGCGGCCCG ACGACGTCTT CATCATTGAC GATGCGGACG AGATCCCGGC CCGTGACGGC  1140
GTCCTTTTCC TCAAGCTCTA CGATGGCTGG ACCGAGCCCT TCGCCTTCCA CATGCGCACG  1200
```

```
TGC CTC ATC TCC TTC CTG CAC TTC TTC AAG ACC CTG TCC TAT GTC      90
Cys Leu Ile Ser Phe Leu His Phe Phe Lys Thr Leu Ser Tyr Val
                20                  25                  30
ACC TTC CCC CGA GAA CTG GCC TCC CTC AGC CCT AAC CTG GTG TCC     135
Thr Phe Pro Arg Glu Leu Ala Ser Leu Ser Pro Asn Leu Val Ser
                35                  40                  45
AGC TTT TTC TGG AAC AAT GCC CCG GTC ACG CCC CAG GCC AGC CCC     180
Ser Phe Phe Trp Asn Asn Ala Pro Val Thr Pro Gln Ala Ser Pro
                50                  55                  60
GAG CCA GGA GGC CCT GAC CTG CTG CGT ACC CCA CTC TAC TCC CAC     225
Glu Pro Gly Gly Pro Asp Leu Leu Arg Thr Pro Leu Tyr Ser His
                65                  70                  75
TCG CCC CTG CTG CAG CCG CTG CCG CCC AGC AAG GCG GCC GAG GAG     270
Ser Pro Leu Leu Gln Pro Leu Pro Pro Ser Lys Ala Ala Glu Glu
                80                  85                  90
CTC CAC CGG GTG GAC TTG GTG CTG CCC GAG GAC ACC ACC GAG TAT     315
Leu His Arg Val Asp Leu Val Leu Pro Glu Asp Thr Thr Glu Tyr
                95                 100                 105
TTC GTG CGC ACC AAG GCC GGC GGC GTC TGC TTC AAA CCC GGC ACC     360
Phe Val Arg Thr Lys Ala Gly Gly Val Cys Phe Lys Pro Gly Thr
               110                 115                 120
AAG ATG CTG GAG AGG CCG CCC CCG GGA CGG CCG GAG GAG AAG CCT     405
Lys Met Leu Glu Arg Pro Pro Pro Gly Arg Pro Glu Glu Lys Pro
               125                 130                 135
GAG GGG GCC AAC GGC TCC TCG GCC CGG CGG CCA CCC CGG TAC CTC     450
Glu Gly Ala Asn Gly Ser Ser Ala Arg Arg Pro Pro Arg Tyr Leu
               140                 145                 150
CTG AGC GCC CGG GAG CGC ACG GGG GGC CGA GGC GCC CGG CGC AAG     495
Leu Ser Ala Arg Glu Arg Thr Gly Gly Arg Gly Ala Arg Arg Lys
```

```
Arg Leu Arg Asn Leu Arg Pro Asp Asp Val Phe Ile Ile Asp Asp
                305                 310                 315

GCG GAC GAG ATC CCG GCC CGT GAC GGC GTC CTT TTC CTC AAG CTC     990
Ala Asp Glu Ile Pro Ala Arg Asp Gly Val Leu Phe Leu Lys Leu
                320                 325                 330

TAC GAT GGC TGG ACC GAG CCC TTC GCC TTC CAC ATG CGC ACG TCG     1035
Tyr Asp Gly Trp Thr Glu Pro Phe Ala Phe His Met Arg Thr Ser
                335                 340                 345

CTC TAC GGC TTC TTC TGG AAG CAG CCG GGC ACC CTG GAG GTG GTG     1080
Leu Tyr Gly Phe Phe Trp Lys Gln Pro Gly Thr Leu Glu Val Val
                350                 355                 360

TCA GGC TGC ACG GTG GAC ATG CTG CAG GCA GTG TAT GGG CTG GAC     1125
Ser Gly Cys Thr Val Asp Met Leu Gln Ala Val Tyr Gly Leu Asp
                365                 370                 375

GGC ATC CGC CTG CGC CGC CGC CAG TAC TAC ACC ATG CCC AAC TTC     1170
Gly Ile Arg Leu Arg Arg Arg Gln Tyr Tyr Thr Met Pro Asn Phe
                380                 385                 390

AGA CAG TAT GAG AAC CGC ACC GGC CAC ATC CTG GTG CAG TGG TCG     1215
Arg Gln Tyr Glu Asn Arg Thr Gly His Ile Leu Val Gln Trp Ser
                395                 400                 405

CTG GGC AGC CCC CTG CAC TTC GCC GGC TGG CAC TGC TCC TGG TGC     1260
Leu Gly Ser Pro Leu His Phe Ala Gly Trp His Cys Ser Trp Cys
                410                 415                 420

TTC ACG CCC GAG GGC ATC TAC TTC AAG CTC GTG TCC GCC CAG AAT     1305
Phe Thr Pro Glu Gly Ile Tyr Phe Lys Leu Val Ser Ala Gln Asn
                425                 430                 435

GGC GAC TTC CCA CGC TGG GGT GAC TAC GAG GAC AAG CGG GAC CTG     1350
Gly Asp Phe Pro Arg Trp Gly Asp Tyr Glu Asp Lys Arg Asp Leu
                440                 445                 450
```

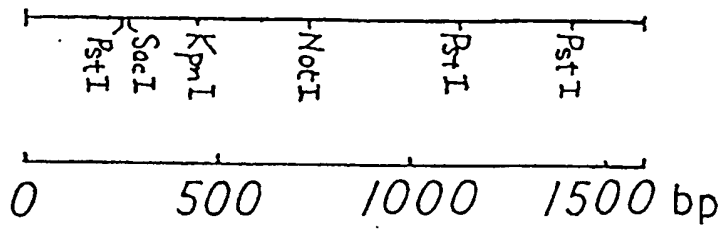

Figure 1

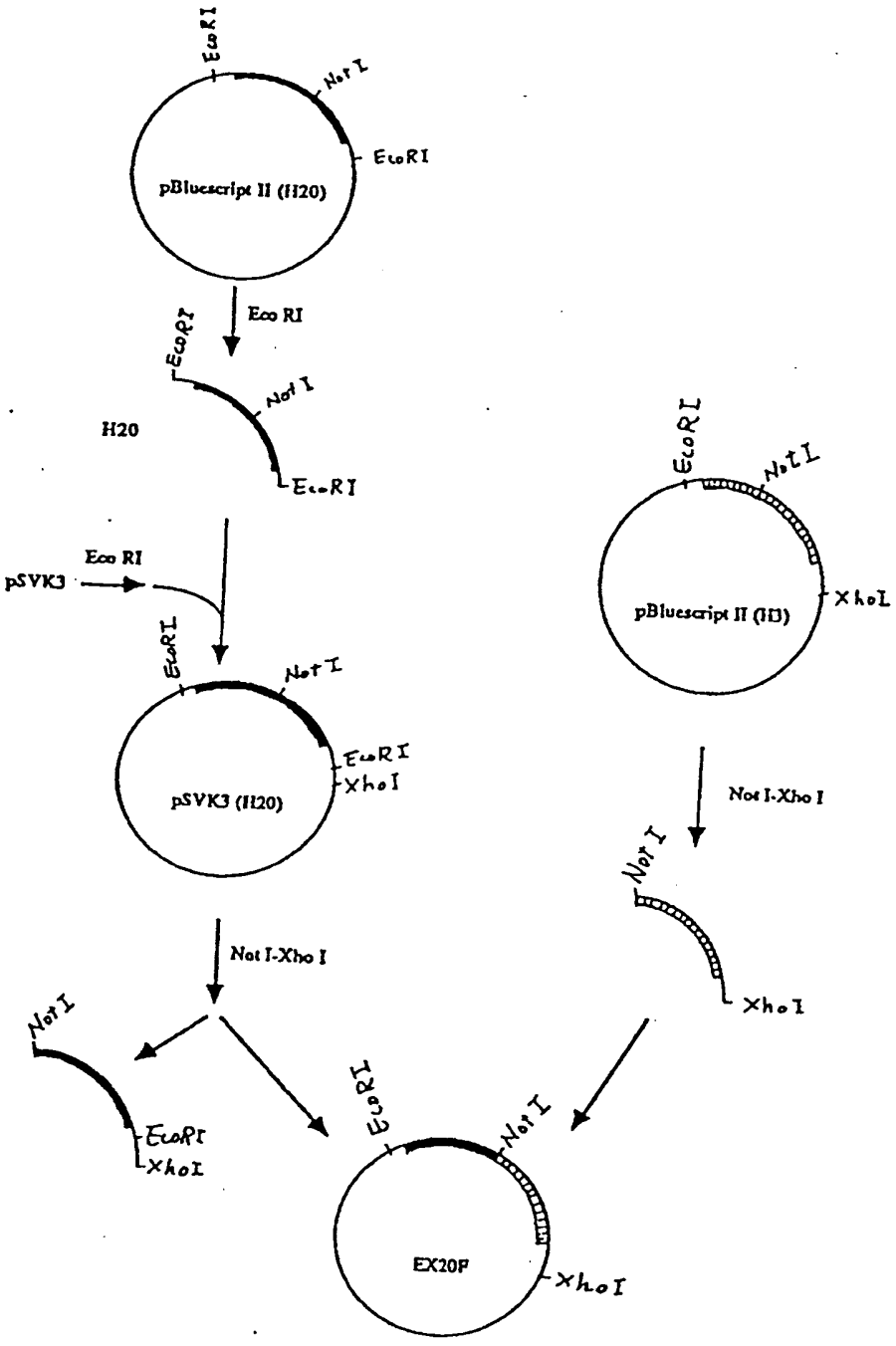
EcoRI
Not I
EcoRI
pBluescript II (H20)
Eco RI
EcoRI
Not I
H20
EcoRI
Eco RI
pSVK3
EcoRI
Not I
EcoRI
XhoI
pSVK3 (H20)
Not I-Xho I
Not I
EcoRI
XhoI
EcoRI
Not I
XhoI
EX20P
EcoRI
Not I
XhoL
pBluescript II (H3)
Not I-Xho I
Not I
XhoI

Figure 3